BLOCK−BASED SPEECH BANDWIDTH EXTENSION SYSTEM WITH SEPERATED ENVELOPE ENERGY RATIO ESTIMATION (WedAmOR6)

Author(s) :

Sheng Yao (City University of Hong Kong, Hong Kong)
Cheung−Fat Chan (City University of Hong Kong, Hong Kong)

Abstract :

The major issue in extending bandwidth of narrowband speech signal (0−4kHz) is the estimation of high−band portion (4−8 kHz) of spectral envelope. It is found that, apart from the shape of high−band spectral envelope, the relative energy level of the missing high band to the observable low band is also crucial to the system performance. In this paper, the two−fold problem is solved by two different estimation rules. In memoryless bandwidth extension systems, the missing high−band information is estimated from narrowband speech using the current frame only. As the narrowband−to−wideband mapping is a one−to−many problem ([1]), memoryless system is likely to cause hissing and whistling artifacts. Our method treats envelope shape estimation on a block basis. Detected narrowband speech block is either one word or a sequence of words, which is modeled by CDHMM (continuous density hidden Markov model) and mapped to a wideband CDHMM pre−trained by original version of the speech block. High−band energy level, present as normalized energy ratio to observable low−band energy, is estimated on an MMSE rule. Both subjective and objective evaluations show that hissing and whistling artifacts are reduced and the spectrally extended wideband speech (0−8kHz) is pleasant to listen.

Menu